Intelligent Microwave Staring Correlated Imaging

Kui Ying, Xinyu Yu, Jiana Shen, Shilu Zhang, and Yuanyue Guo[*]

Abstract—Microwave staring correlated imaging (MSCI) is a super-resolution imaging technique based on temporal-spatial stochastic radiation fields (TSSRFs), which requires an accurate calculation of the electromagnetic field at the imaging plane. However, systematic errors always exist in practice, such as the time synchronization and frequency synchronization errors of radar systems, which make it difficult to calculate the required TSSRFs accurately, and this deteriorates the imaging results. Meanwhile, some imaging algorithms have problems such as high computational complexity. In this paper, an intelligent MSCI method based on the deep neural network (DNN) is proposed, which can accomplish imaging directly from the echoes, avoiding the computation of TSSRFs. A multi-level residual convolutional neural network (MRCNN) is developed for the DNN, and simulations and experiments are carried out to obtain the dataset for training and testing the MRCNN. Compared with the conventional MSCI methods, the imaging results verify the effectiveness of intelligent MSCI in terms of imaging quality and computational efficiency.

1. INTRODUCTION

Microwave imaging can be accomplished by acquiring target electromagnetic scattering information in the microwave frequency, which has all-weather and all-time imaging capability for broad applications in remote sensing, military target investigation. Real aperture radar (RAR) imaging allows continuous observation of the focal area, but its azimuthal resolution is limited by its antenna aperture size, so it can only be applied to low-resolution imaging [1, 2]. Synthetic aperture radar (SAR) and inverse synthetic aperture radar (ISAR) imaging rely on the relative motion between radar and target to synthesize large antenna apertures to achieve high-resolution imaging, so it is not achievable to staring imaging for a long time at the observation scene [3, 4]. Microwave staring correlated imaging (MSCI) is a new imaging regime that allows super-resolution staring imaging beyond the limits of its antenna aperture size, and its central idea is that the random radiation source (RRS) generates temporal-spatial stochastic radiation fields (TSSRFs) to modulate different target information within its antenna beam, so the echo signals contain identifiable target information, and by correlating the echoes with the preset radiation fields, target information can be decoupled to obtain high-resolution target imaging [5, 6].

The imaging quality of MCSI is influenced by the computational accuracy of the TSSRFs as well as the correlated reconstruction algorithm. On the one hand, there are always systematic errors in real MSCI, such as the position errors of the transmitting and receiving arrays, the time synchronization and frequency synchronization errors of radar systems, etc., which seriously affect the accurate calculation of the preset TSSRFs and result in a serious mismatch with the scattered echo signal collected at the corresponding time, degrading the performance of the MSCI [7–14]. In recent years, researchers have proposed some solutions for each different influencing factor. The array position errors are studied in MSCI, and the basis pursuit (BP) algorithm is proposed by iterating alternately between the imaging and

the error estimation to obtain better imaging results [7]. Afterward, the self-calibration methods based on orthogonal match pursuit (OMP) and sparse Bayesian learning (SBL) are proposed to reconstruct the target images and estimate the amplitude-phase errors of the signals from different transmitting antennas [8, 9]. Besides, the errors caused by random amplitude, phase, and synchronization exist in each transmitted pulse of each channel, Cao et al. regarded it as additive perturbations to the radiation field and proposed an iterative optimization method based on FOCUSS and TV-TLS, which estimated the perturbation and scattering coefficient of the target to obtain better imaging quality respectively [10, 11]. Overall, the methods above address only one of the factors; however, it is difficult to obtain all the system parameters accurately to calculate TSSRFs in engineering. On the other hand, in the MSCI reconstruction algorithm, He et al. and Meng et al. introduced regularization methods to MSCI and proposed Tikhonov regularization, truncated singular value decomposition (TSVD) method, total variance regularization, and iterative regularized inversion algorithms based on Landweber's algorithm, which enhances the stability of the problem solution by limiting the smaller singular value components of the observation matrix based on the least-squares method [12, 13]. For sparse targets, OMP, SBL, and other sparse reconstruction methods are applied to MSCI and achieve better imaging results [14]. These imaging optimization algorithms still suffer from the high computational complexity and long imaging times.

In recent years, owing to the easily accessible large amount of data and growing computational power, deep learning has demonstrated excellent performance in several areas of regression and classification. Constructing specific neural network structures has good applications in different fields, such as image classification and recognition using convolutional neural networks (CNNs) [15, 16], generation of the desired high-quality images using generating adversarial networks (GANs) [17, 18], and the prediction of dense data using full convolutional networks (FCNs) [19–21]. Recently, neural networks have been applied in microwave imaging. In parameter prediction of SAR target simulation, Qiu's team used CNN to establish the mapping relationship between simulation images and simulation parameters, and reconstructed the images from the predicted parameters by neural networks, which have better performance than images reconstructed from manually set parameters, and realized the inverse process of inferring simulation parameters from real images [22]. In nonlinear electromagnetic inverse scattering imaging, Shao and Du developed an auto-encoder to find compressed alternative solutions for high-resolution target images, and then used a deep learning network (DLN) to convert the electromagnetic signal into a compressed alternative solution, and subsequently decompressed it to the target image, which realized the process of reconstructing a nonlinear electromagnetic signal into a target medium image [23]. In intelligent electromagnetic perception, Li et al. combined a programmable metasurface with adaptively modulated electromagnetic waves and the artificial neural network IM-CNN-1, and then reconstructed images of the human body from measured microwave data and identified the region of the hand and chest area within the whole image using Faster R-CNN, controlling the metasurface by the optimal digital coding pattern to focus its radiation wave onto the desired spots and reading hand signals and physiological state by the artificial neural network IM-CNN-2 [24, 25].

MSCI is a linear electromagnetic inverse scattering process that correlates the scattered echoes of targets with the preset TSSRFs to reconstruct the target images, where the mapping relationship between the scattered echoes and the target imaging is closely related to linear regression in deep learning [19, 21, 26]. In this paper, a fundamental connection between the deep neural network (DNN) architecture and MSCI is first proposed. Inspired by this connection, we develop a multi-level residual convolutional neural network framework for MSCI called MRCNN. In the MSCI system, MRCNN builds an intelligent correlated imaging mapping model, which constructs an end-to-end mapping between the scattering echo and the target imaging after autonomous learning. Next, a numerical simulation platform is set up, where TSSRFs are generated in the target region by emitting amplitude-hopping and phase-hopping waveforms from multi-transmitter systems, and 10,000 mixed National Institute of Standards and Technology database (MNIST) handwritten digital datasets [27] are binarized to the target's backscattering coefficient, with the received scattered echoes in the simulated MSCI system as the output and input of the MRCNN, respectively. Subsequently, an experiment is carried out in an anechoic chamber. By constructing a set of orthogonal target scenes consisting of corner reflectors, a metamaterial aperture antenna serves as the RRS, and different TSSRFs are generated by frequency sweeping that radiate into the imaging area, and the received scattered echoes with the distribution of

the backscattering coefficients in the imaging region are used as the input and output of the MRCNN, respectively. The samples are randomly grouped for training and testing MRCNN.

Simulated and experimental results show that MRCNN demonstrates the ability to image unknown targets after autonomous learning, which proves the feasibility of MRCNN. Since it uses convolution to extract high-dimensional information directly from the echoes and deconvolution to map the high-dimensional feature information to the target image, this intelligent MSCI avoids the calculation of TSSRFs with RRS parameters and the use of complex and time-consuming iterative optimization algorithms, providing a realistic solution for the application of MSCI.

2. PROBLEM STATEMENT

This section reveals the connection between the DNN structure and MSCI. Since in an MSCI system, there is a strong linear end-to-end mapping between the scattered echoes obtained in the different TSSRF modes and the target imaging, the DNN can learn end-to-end predictions of complex data because of its linear regression capability, which suggests that DNN can provide an efficient alternative solution in MSCI.

2.1. Connection between DNN Architecture and MSCI

The diagram of a typical MSCI system with multiple transmitters and one receiver is illustrated in Figure 1. The RSS consists of N antennas distributed randomly in the XOY plane. The center of the antenna array is the origin O, and the position of the ith antenna is $\vec{r_i}$, $i = 1, 2, \ldots, N$. The TSSRFs are generated in the target region by emitting amplitude-hopping and phase-hopping waveforms from multi-transmitter systems, where the TSSRFs generated at the mth radiation are denoted as E_{inc}^m. The 2-D imaging area S is located at the upper right of the RSS; its height is H; and the horizontal distance from the origin is L_1. After the TSSRFs continue to propagate, interacting with targets in the imaging area to produce scattered fields, the scattered echo signal is received by the receiving antenna. The horizontal distance from the receiving antenna to the origin is L_2, which is located in the XOY plane.

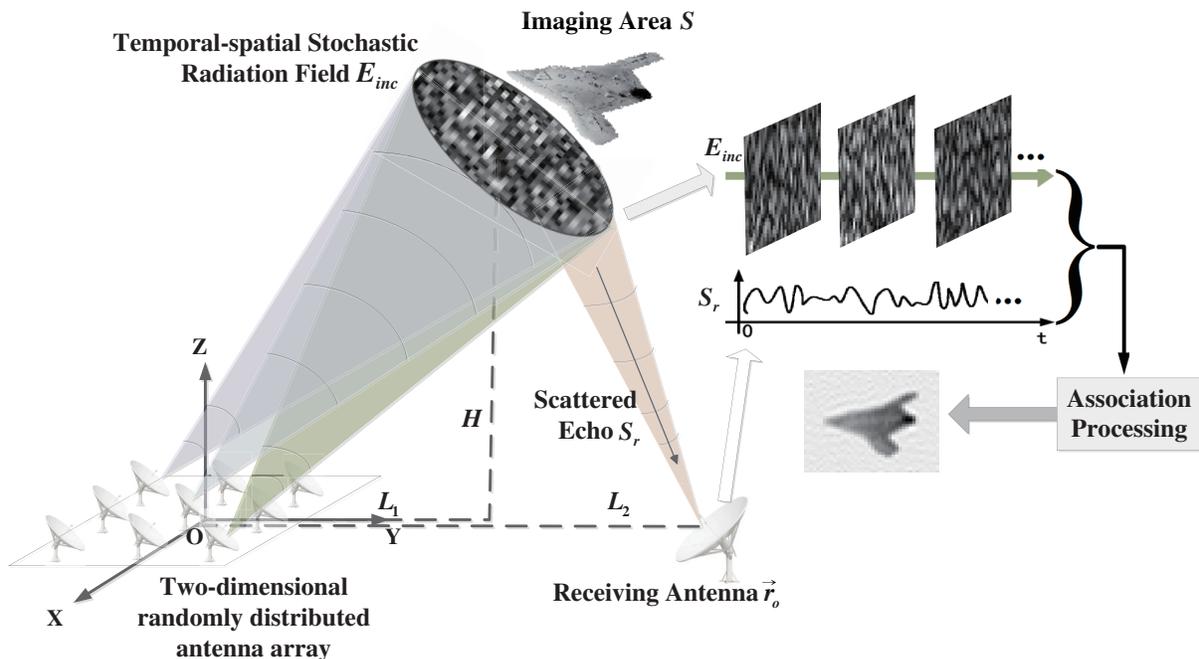Assume that the orthogonal, independent signal $S_i(\vec{r_i}, t)$ emitted by the ith transmitting antenna

Figure 1. Microwave correlated imaging system.

can be expressed as:

$$S_i\left(\vec{r_i}, t\right) = \sum_{q=1}^{Q} s_{i,q}(t - qT)\, \text{rect}\left[\frac{t - (q-1)T}{\tau}\right] \tag{1}$$

where Q is the total number of pulses emitted in one imaging process, $s_{i,q}(t)$ the qth pulse signal emitted by the ith transmitting antenna, T the pulse period, τ the pulse width, and $T \gg \tau$, rect[] the rectcorner window function.

The TSSRFs $E_{inc}(\vec{r}, t)$ constructed by the N transmitters in the target area can be expressed as:

$$E_{inc}(\vec{r}, t) = \sum_{i=1}^{N} \frac{F_i\left(\vec{r_i}, \vec{r}\right)}{4\pi\left|\vec{r} - \vec{r_i}\right|} S_i\left(\vec{r_i}, t - \frac{\left|\vec{r} - \vec{r_i}\right|}{c}\right) \tag{2}$$

where $\vec{r_i}$ is the position vector of the ith antenna, $\vec{r}$ the position vector of the target, $F_i\left(\vec{r_i}, \vec{r}\right)$ the antenna pattern of the ith antenna, and c the speed of light.

After the target reflection and second propagation in free space, the received echo signal $S\left(\vec{r_o}, t\right)$ can be expressed according to the first-order Bohr approximation as:

$$S_r\left(\vec{r_o}, t\right) = \iint_S \sum_{i=1}^{N} \frac{F_i(\vec{r_i}, \vec{r}) F_o\left(\vec{r_o}, \vec{r}\right)}{(4\pi)^2 \left|\vec{r} - \vec{r_i}\right| \left|\vec{r} - \vec{r_o}\right|} S_i\left(\vec{r_j}, t - \frac{\left|\vec{r} - \vec{r_i}\right| + \left|\vec{r} - \vec{r_o}\right|}{c}\right) \sigma\left(\vec{r}\right) d\vec{r} + n(t) \tag{3}$$

where $\vec{r_o}$ is the position vector of the receiving antenna, $F_o\left(\vec{r_o}, \vec{r}\right)$ the antenna pattern of the receiving antenna, $\sigma(\vec{r})$ the backscattering coefficient of the target at $\vec{r}$, and $n(t)$ the noise.

Considering the propagation time delay from the target to the receiver, the modified TSSRF is described as:

$$E_{rad}\left(t, \vec{r}\right) = \sum_{i=1}^{N} \frac{F_i\left(\vec{r_i}, \vec{r}\right) F_o\left(\vec{r_o}, \vec{r}\right)}{(4\pi)^2 \left|\vec{r} - \vec{r_i}\right| \left|\vec{r} - \vec{r_o}\right|} S_i\left(\vec{r_j}, t - \frac{\left|\vec{r} - \vec{r_i}\right| + \left|\vec{r} - \vec{r_o}\right|}{c}\right) \tag{4}$$

The relationship between the received echo $S\left(\vec{r_o}, t\right)$ and the TSSRF $E_{rad}\left(t, \vec{r}\right)$ can be described as:

$$S_r\left(\vec{r_o}, t\right) = \iint_S E_{rad}\left(t, \vec{r}\right) \sigma\left(\vec{r}\right) d\vec{r} + n(t) \tag{5}$$

In data processing, the received echo signals are a set of discrete-time samples, so the reconstruction process of the target must be discretized in both the spatial and temporal domains. The imaging area is divided into Q discrete grids, and the observation time is divided into M samples. Then the discrete form of Equation (5) is:

$$\begin{bmatrix} S_r\left(t_1\right) \\ S_r\left(t_2\right) \\ \vdots \\ S_r\left(t_M\right) \end{bmatrix} = \begin{bmatrix} E_{rad}\left(t_1, \vec{r_1}\right) & E_{rad}\left(t_1, \vec{r_2}\right) & \cdots E_{rad}\left(t_1, \vec{r_Q}\right) \\ E_{rad}\left(t_2, \vec{r_1}\right) & E_{rad}\left(t_2, \vec{r_2}\right) & \cdots E_{rad}\left(t_2, \vec{r_Q}\right) \\ \vdots & \vdots & \ddots \\ E_{rad}\left(t_M, \vec{r_1}\right) & E_{rad}\left(t_M, \vec{r_2}\right) & \cdots E_{rad}\left(t_M, \vec{r_Q}\right) \end{bmatrix} \begin{bmatrix} \sigma\left(\vec{r_1}\right) \\ \sigma\left(\vec{r_2}\right) \\ \vdots \\ \sigma\left(\vec{r_Q}\right) \end{bmatrix} + \begin{bmatrix} n\left(\vec{r_1}\right) \\ n\left(\vec{r_2}\right) \\ \vdots \\ n\left(\vec{r_Q}\right) \end{bmatrix} \tag{6}$$

Its vector form is:

$$\vec{S}_r = \vec{E}_{rad} \cdot \vec{\sigma} + \vec{n}_0 \tag{7}$$

where $\vec{S}_r$ is the sampling echo signal at the receiver using the Analog to Digital Converter (ADC) device, $\vec{E}_{rad}$ the TSSRF matrix vector, $\vec{n}_0$ the noise vector, and $\vec{\sigma}$ the observed target backscattering coefficient vector.

Based on the scattered echo vector $\vec{S}_r$ and TSSRF vector $\vec{E}_{rad}$, the reconstructed images $\vec{\hat{\sigma}}$ can be described as:

$$\vec{\hat{\sigma}} = \xi\left[\vec{E}_{rad}, \vec{S}_r\right] \tag{8}$$

where ξ denotes the correlated processing operator.

The common correlated imaging algorithms can be applied to solve (8), including pseudo-inverse, matched filter method, etc., and the accurate calculation of the TSSRFs is necessary for imaging. From (4), the parameters of the RRS must be known for the calculation of the TSSRFs; however, in practical

imaging, it is difficult to synchronize time and frequency precisely in radar systems to calculate TSSRFs accurately. When higher quality imaging needs are required, further optimization algorithms must be used, such as TSVD regularization, Tikhonov regularization, and sparse Bayesian learning, which is a complex and time-consuming iterative problem of convex optimization.
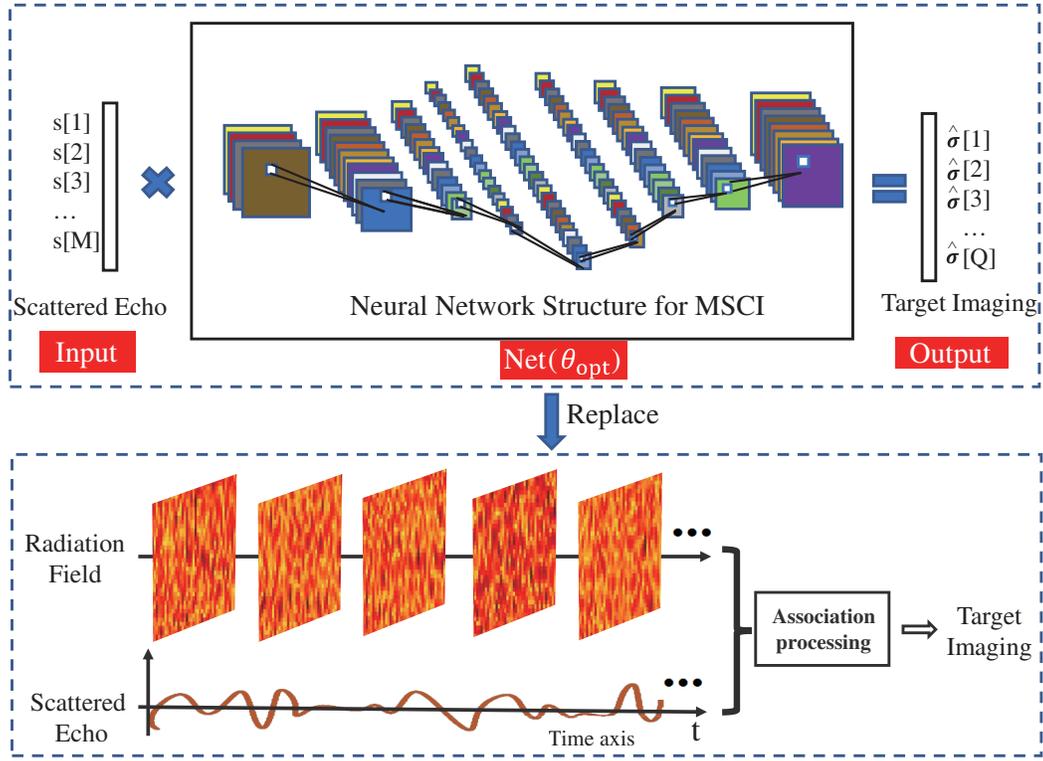
Figure 2. DNN mapping model. It constructs an end-to-end mapping between scattered echoes and target imaging.

For a deep learning neural network with linear regression analysis, the essence is exactly the linear regression function $\vec{Y} = \vec{W}\vec{X} + \vec{B}$, which is linearly transformed from one feature space $\vec{X}$ to another $\vec{Y}$, and $\vec{W}$ and $\vec{B}$ are the weight and bias, respectively. Inspired by this observation, an intelligent MSCI method based on the DNN is proposed to build, as shown in Figure 2, which constructs an end-to-end mapping between the scattering echo $\vec{S}_r$ and the target imaging $\vec{\hat{\sigma}}$ after autonomous learning, described as:

$$\vec{\hat{\sigma}} = \text{Net}\left(\vec{S}_r, \theta\right) \tag{9}$$

where Net represents the DNN, and $\theta = \{\vec{W}, \vec{B}\}$ is network parameters of the neural network under the MSCI system, respectively. The back propagation updates the parameters with a gradient descent optimization algorithm and a mean square error loss function, and the process is represented as:

$$\theta^{k+1} = \theta^k - \eta \frac{\partial \left[\frac{1}{p}\sum_{p=1}^{P}\left(\text{Net}\left(\vec{S}_r^{\,p}, \theta^k\right) - \vec{\sigma}^p\right)^2\right]}{\partial \theta} \tag{10}$$

where η is the learning rate; k is the iterative; $\vec{S}_r^{\,p}$ and $\vec{\sigma}^p$ are the scattered echoes and target backscattering coefficient in the pth target scene set, respectively; P is the number of target scene sets for each iteration of training. As shown in Eq. (11), the optimal parameters θ_{opt} are obtained after

autonomous learning, and then the scattered echoes of the unknown target scene are then input to the DNN, which directly outputs its target image $\vec{\hat{\sigma}}$:

$$\vec{\hat{\sigma}} = \text{Net}\left(\vec{S}_r, \theta_{opt}\right) \tag{11}$$

When a set of available samples is acquired, it is permissible to train both $\vec{W}$ and $\vec{B}$ at each layer of the DNN to find the optimal network parameter model. Comparing this approach to conventional MSCI methods, the learned method would be more efficient as it optimizes the weighting matrices and biases and the reconstruction error with respect to the target backscattering coefficient. In general, the above observations suggest that DNNs are well suitable for MSCI problems.

2.2. Multi-Level Residual Convolutional Neural Network Structure for MSCI

After demonstrating the connection between DNN architecture and MSCI, we now design a multi-level residual convolutional neural network structure to solve the MSCI problem. The idea of convolution and deconvolution is proposed in semantic segmentation following the FCN proposed by Long et al. [19]. The idea of the MRCNN is constructed by using the strong feature extraction capability of the convolutional layer network and the target feature mapping imaging capability of the deconvolutional layer network, where the convolutional layer network finds the high-dimensional feature information from the scattered echo, and then the deconvolutional layer network maps the high-dimensional feature information to the target image. Since a strong linear link is established between the received scattered echoes and target imaging, a large number of convolutional and deconvolutional layers are applied to increase the depth of the network; however, this increases the training difficulty of the network. To solve the gradient explosion and gradient disappearance problems caused by increasing network depth, both residual fusions in the convolutional and deconvolutional layers are invoked following the residual network (Resnet) proposed by He et al. [28], so that more complex extraction of feature patterns can be performed without causing the network overfitting. In addition, in the network construction, to effectively retain the target feature information in the scattering echo for imaging, the convolution layer is not normalized using the batch normalization (BN) layer, and to stabilize the network structure, the deconvolution layer is normalized using the BN layer.

The scheme of the present network is shown in Figure 3, and the MSCI system presented in Figure 1 is simulated in our numerical simulation platform. By controlling the phase and amplitude of the emitted signals from the 24 transmitting antennas in the RRS, 160 groups of 160 TSSRFs are generated in the target area, with each group of TSSRFs operating at 160 frequencies within the operating bandwidth of the antenna array, obtaining 160×160 scattered echoes from multiple targets as input to the MRCNN. Then it is transformed into tensors of $40 \times 40 \times 64$ after using two convolution kernels of 7×7. Next, eight convolutional residual modules are transformed into high-dimensional features of $5 \times 5 \times 512$. Each module uses two convolutional kernels of 3×3, and the output tensors of each two modules are $20 \times 20 \times 128$, $10 \times 10 \times 256$, and $5 \times 5 \times 512$, respectively. The ReLU function is chosen as the activation function for the deconvolution layer to extract high-dimensional features. After obtaining $5 \times 5 \times 512$ high-dimensional feature information by convolutional layer network, it is then converted into tensors of $10 \times 10 \times 256$ by a convolutional kernel of 3×3 with a step of 2 and a padding of 2 in the deconvolution layer. The next is three deconvolution residual modules with the same settings, which convert it into tensors of $80 \times 80 \times 8$, and the output tensors of each residual module are $20 \times 20 \times 128$, $40 \times 40 \times 64$, $80 \times 80 \times 32$, respectively. The Softplus function is chosen as the activation function for the deconvolution layer to reconstruct the target image. Then the tensors with the same size as the input are obtained by one layer of deconvolution, and the feature numbers are changed to 2 by a $1 \times 1 \times 2$ convolution. Finally, compared with the binarized microwave image of the corresponding target, the binary cross-entropy is used as the loss function, and a quantitative gradient descent optimization algorithm is driven to update the parameters by back propagation. In the following section, we will verify the applicability of MRCNN for MSCI in simulation and experiment.
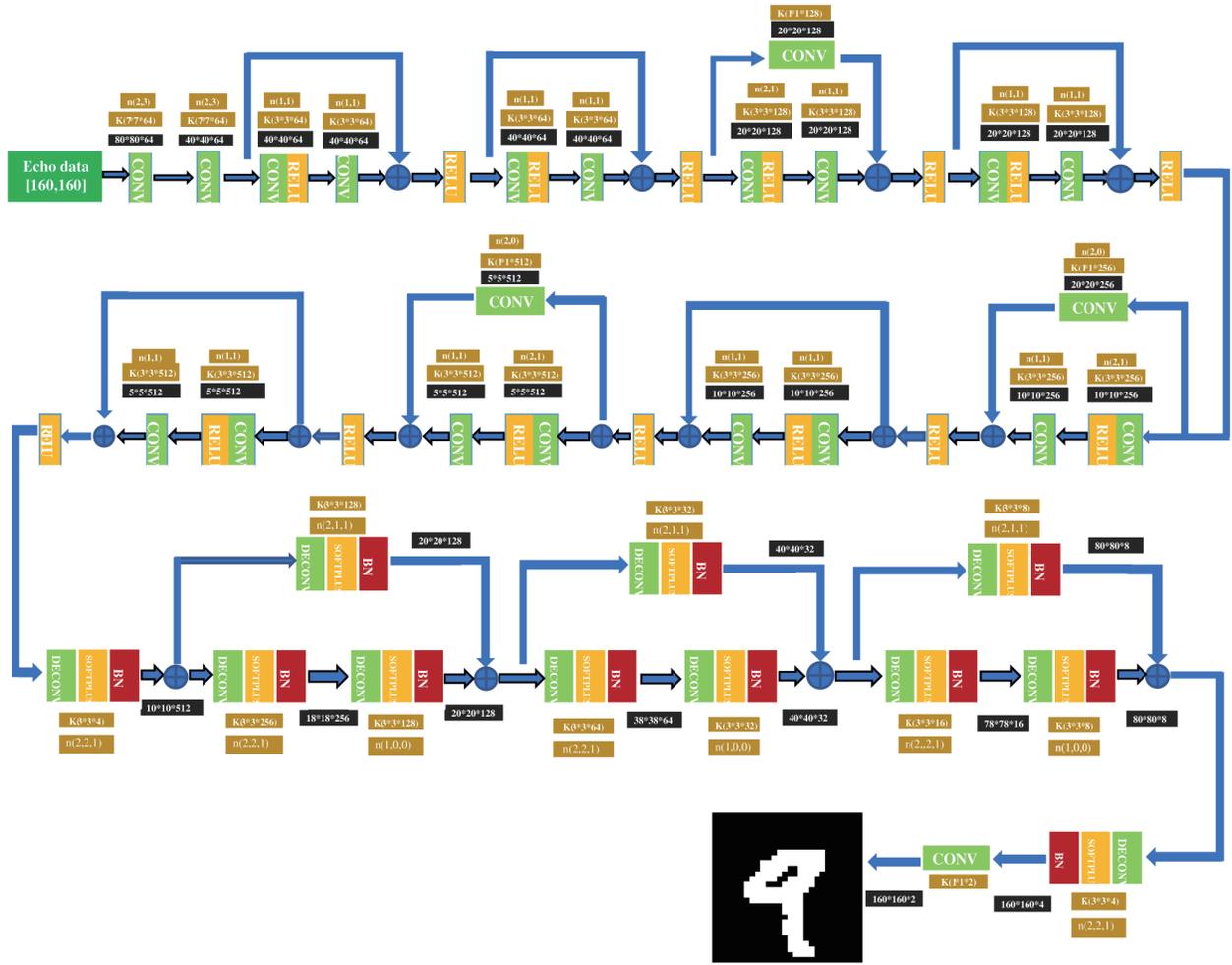
Figure 3. Multi-level residual convolutional neural network structure (i.e., MRCNN). BN is the batch normalization, SOFTPLUS, RELU is the nonlinear activation function, $k(a, b, c)$ is the convolution kernel of size $a \times b \times c$, and $n(a, b)$ represents stride $= a$ and padding $= b$, $a \times b \times c$ are the out tensors of this layer.

3. NUMERICAL SIMULATION AND RESULT ANALYSIS

In this section, the performance of MRCNN in solving MSCI problems is demonstrated through numerical simulations. For comparison, we report the corresponding results of the generalized minimum residual (GMRES) optimization algorithm to reconstruct the target image.

## 3.1. Establishment of Numerical Simulation Platform

The simulated imaging scenario is shown in Figure 1. The RRS consists of 24 randomly distributed antennas located in the XOY plane, and the center of the array is the origin O. The positions of the 24 antennas are shown in the red rectangles in Figure 4. The TSSRFs are generated by controlling the phase and amplitude of the emitted signal of each transmitting antenna. The random amplitude of each transmitter's signal is one of $[0, 0.2, 0.4, 0.6, 0.8, 1]$, and the random phase is one of $[0, \pi]$, generating 160 groups of TSSRFs in the imaging area, each group with 160 different TSSRFs. The radiated 2-D imaging region S is located at the upper right of the antenna array, whose height H is 128 m. The imaging region S of 20 m × 20 m is divided into spatial grids of 160×160, with the length of each resolution cell of 0.125 m. The receiving antenna is located in the XOY plane with a horizontal distance to the origin of
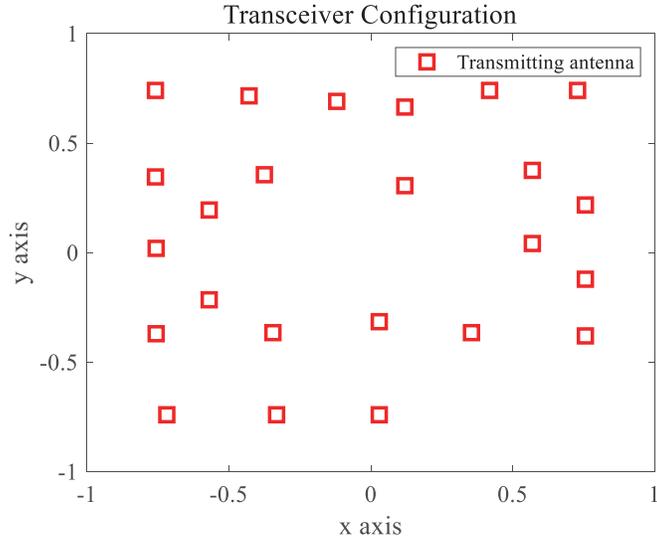
Figure 4. Position of each transmitting antenna of the microwave RRS.

100 m. 10,000 MNIST datasets are binarized into images into 160×160 images, a handwritten numeric database containing 0 to 9, simulating the target with a scattering coefficient of 1 and the background with a scattering coefficient of 0 on different resolution cells in the imaging region. The TSSRFs interact with the scattering coefficient, and then the scattered echoes are received by a single receiver.

According to the system scenario, the operating bandwidth of the system is set between 9.5 GHz and 14 GHz, and 160 TSSRFs in each group work at 160 interval frequencies within the operating bandwidth of the antenna array. The scattered echoes at each TSSRF are obtained by scanning the frequency in steps of 28.125 MHz. Thus, the generated 10,000 sets of scattered echo signals of 160×160
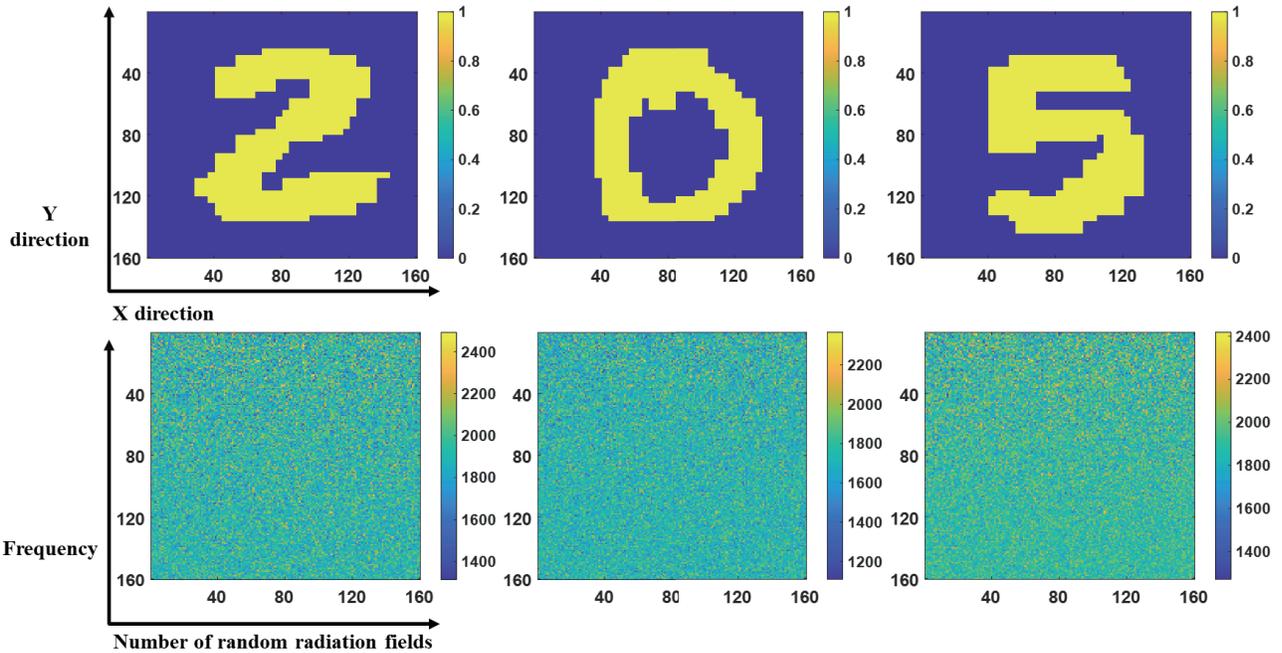
Figure 5. Schematic visualization of the target and echo data. The first row shows the target with a scattering coefficient of 1 and the second row shows the corresponding amplitude of the echo data in size 160×160.

are used as the input to the MRCNN, and the corresponding 10,000 target backscatter coefficients are used as the output of the MRCNN. The calculation of TSSRFs and GMRES algorithm for conventional MSCI are implemented in MATLAB 2021. The target and echo data are shown in Figure 5. As shown in Figure 5, each row of echo data is 160×1 in size and is distributed over each equally spaced frequency within the operating bandwidth of 9.5–14 GHz, with 160 rows corresponding to 160 frequencies. Each row of echo data is obtained by transmitting random amplitude-hopping and phase-hopping signals from a multi-transmitter system to generate 160 random radiation fields, which are then radiated to the target area and scattered to the receiving antenna, containing 160 types of target information. The system parameters are given in Table 1.

Table 1. Simulation parameters.

Parameter	Value	Parameter	Value
Size of antenna array	$2\,\mathrm{m} \times 2\,\mathrm{m}$	Pulse width	100 ns
Antennas Number	24	Pulse of period	1000 ns
Carrier frequency	9 GHz	Target distance	128 m
Frequency bandwith	4.5 GHz	Imaging area	$20\,\mathrm{m} \times 20\,\mathrm{m}$
Number of pulses	1800	Grid size	0.5 m

3.2. Training and Testing of MNIST Dataset

The process of using neural network training for imaging is shown in Figure 6. The process can be divided into two stages. The first stage is the training of the network model: first, in the MSCI system shown in Figure 1, the MNIST dataset is binarized into the backscattering coefficients of the target region as labels, and the simulation parameters of the MSCI system are set; then, the whole process of MSCI is simulated by MATLAB to obtain the scattering echo; finally, the scattering echoes containing information about different targets and corresponding labels are used to train the neural network. Through iterative training, the neural network can learn the end-to-end mapping between the scattered echo and the target imaging, and each set of echoes and labels in the training set optimizes the weights and biases. The second stage is the testing of the network model: the unknown digital target is simulated in the same MSCI simulation scenario to obtain the scattering echo, which is input to the well-trained network to predict the corresponding target imaging by the mapping relationship; then it is compared with the target to evaluate the effectiveness of the method. 10,000 groups of echoes and labels are divided into two parts, with 7,000 groups for training and 3,000 groups for testing.

Table 2. Network parameters.

Hyper-parameter	Specification
Optimizer	gradient descent with momentum
Loss function	binary cross-entropy function
Echo patch size	160×160
Batch size	4
Epoch number	300
Learning rate range	$[10^{-2},\ 10^{-5}]$
Input channels	1
Layers of ResBlock	11
Total Layers	29
Convolution kernel size	$1 \times 1, 3 \times 3, 7 \times 7$
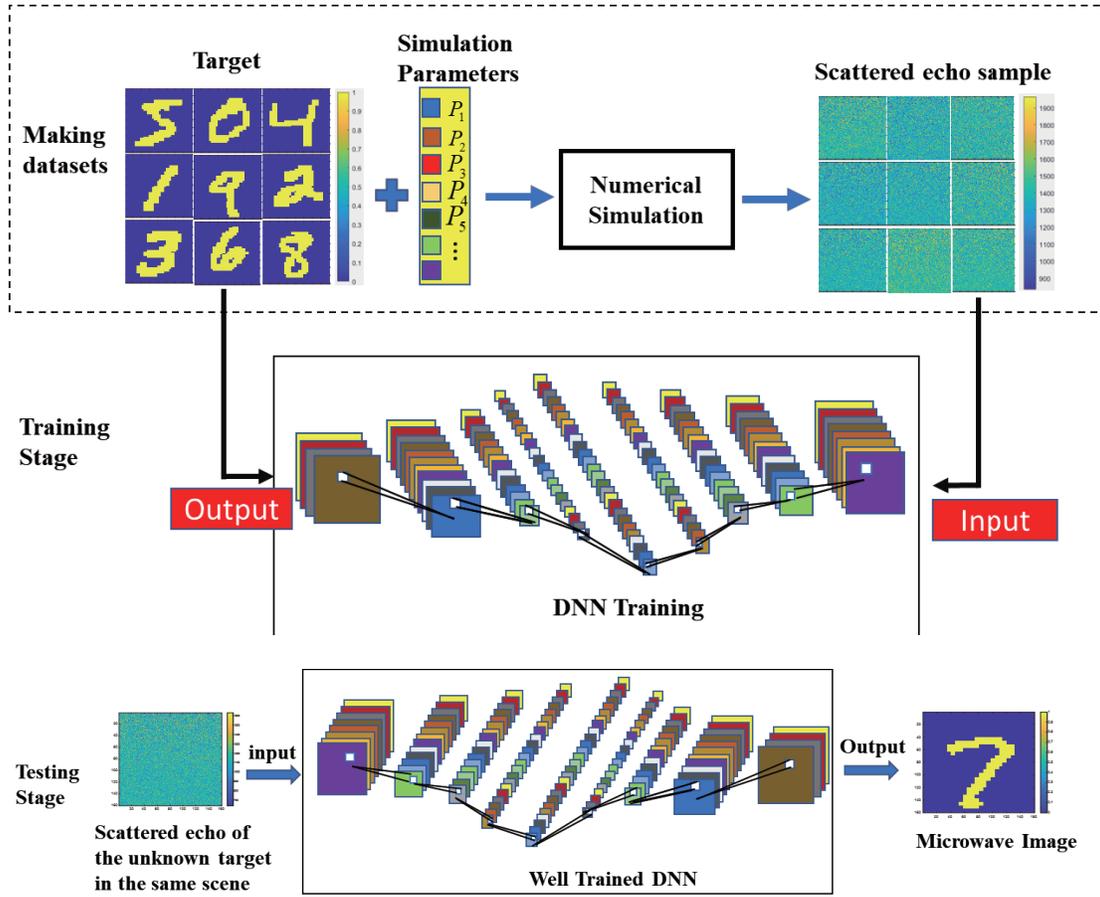
Figure 6. The process of training the neural network and using a well-trained neural network to test imaging.

The structure of the MRCNN used is shown in Figure 3, and the network parameters are shown in Table 2 under the best imaging performance. It is designed with the PyTorch library and takes approximately 8 hours to train. The training stage was performed with a gradient descent optimization algorithm with driving volume, binary cross-entropy as the loss function, a minibatch size of 4, and an epoch setting of 300. The learning rate is set to 10^{-2} and halved after every 50 epochs. The weights and biases are initialized by random weights with the normal distribution of zero means. The calculations are performed in an AMD 5800H 8-core processor and NVIDIA GeForce RTX 3050 Ti server with 128 GB of access memory.

The first and third rows in Figure 7 represent 8 targets, and the second and fourth rows in Figure 7 represent the output images obtained by MRCNN for the corresponding echo data in the training and test sets, respectively. To characterize the recovery image quality, the normalized mean square error (NMSE) is used as a qualitative metric, described as:

$$\text{NMSE} = \frac{\|\hat{\sigma} - \sigma\|_2}{\|\sigma\|_2} \tag{12}$$

where σ and $\hat{\sigma}$ are the scattering coefficient of the target and the reconstructing target images.

Table 3 reports the NMSE results of the reconstructed images in the training and test sets, respectively. Figure 8 shows that the training and test losses of the MRCNN have reached 0.1 during 300 epochs, and the losses become stable which indicates that the MRCNN has achieved a direct mapping link between scattered echoes and target imaging.
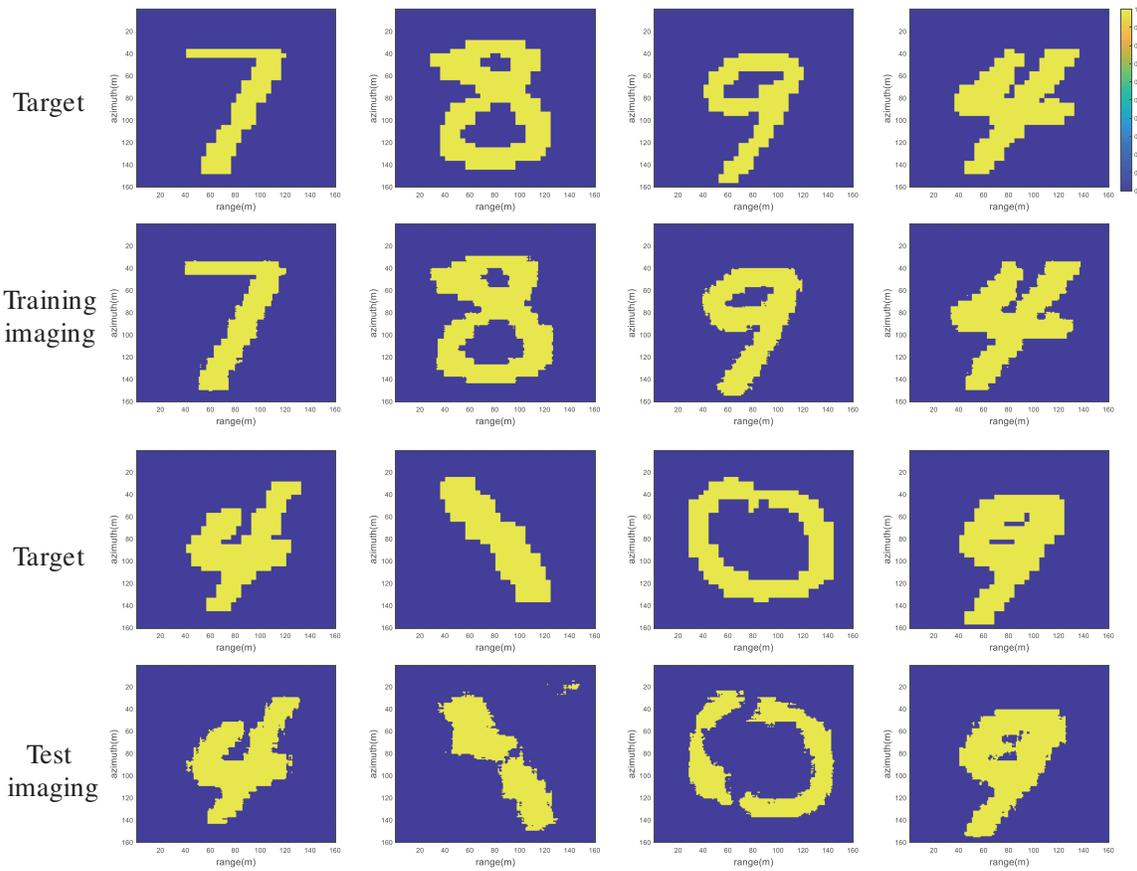
Figure 7. Reconstruction results of MRCNN in training and test sets. The first and third rows show 8 targets, and the second and fourth rows show the output images by MRCNN from the corresponding echo data in the training and test sets, respectively.
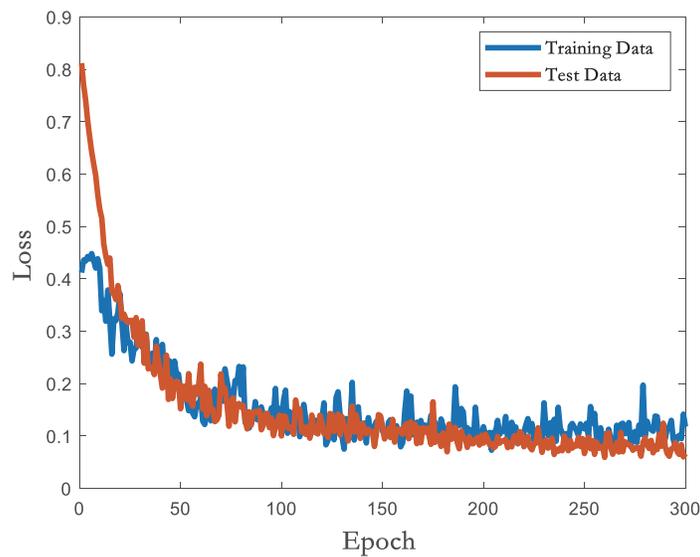
Figure 8. The training loss and test loss of MRCNN for the MNIST data during 300 epochs.

Table 3. NMSE results for the reconstructions in Figure 7.

Column Sequence	1	2	3	4
Training Set	0.0026	0.0035	0.0036	0.0024
Testing Set	0.0042	0.0086	0.0092	0.0083

3.3. Testing Letter Targets with Trained Networks

To further investigate the generalization capability of the MRCNN designed for MSCI, another set of simulations is conducted, where the network is trained with the MNIST dataset, and the test target is the English alphabet, whose backscattering coefficient is 1 on the spatial grid of 160×160. In the MSCI scenario shown in Figure 1, system parameters are all the same as the training data set. The same 160 groups of 160 TSSRFs in random amplitude-phase are obtained in the simulation, interacting with the target's backscattering coefficients, and by scanning 160 frequencies within the operating bandwidth of the antenna array, the scattered echoes of 160×160 in each TSSRF pattern were obtained, which is then input to the well-trained MRCNN.

Figure 9 shows the reconstruction results based on different imaging methods. The first row shows the target, and the second, third, and fourth rows show the imaging results of the backward projection (BP) algorithm, GRRES algorithm, and MRCNN, respectively. To compare the imaging quality of the three imaging methods, the NMSE results corresponding to the different methods are shown in Table 4. As shown in Table 4, the NMSE of the MRCNN reconstructed target images is less than 0.1, so in this case, the real image is very different from the training sample, and the results obtained from the trained MRCNN are satisfactory. It should be noted that MRCNN establishes an intelligent MSCI mapping model, which inputs the scattered echo into the well-trained network and takes less than 1 s to generate the target image through the well-trained network model, and the amount of information required for intelligent MSCI is much less than other algorithms. Among them, both the BP algorithm and the GMRES algorithm take about 10 s. The reconstruction results show that the intelligent MSCI is significantly better than both the BP algorithm and the conventional MSCI optimization algorithm in both imaging quality and efficiency.

**Table 4.** NMSE results for the reconstructions in Figure 10.

Letter Target	U	S	T	C
BP	0.8862	0.8461	0.7659	0.7836
GMRES	0.5263	0.5148	0.5296	0.5489
MRCNN	0.0562	0.0592	0.0583	0.0651

From the above discussion, an important conclusion can be drawn: the intelligent MSCI model based on the MRCNN has been learned in the same MSCI scenario, or the DNN architecture has established a linear electromagnetic inverse scattering network link between the scatter echo and target imaging.

4. EXPERIMENTAL VERIFICATION

4.1. Measurement Setup

To verify the practicality of this intelligent MSCI method, this section presents experiments carried out in a microwave anechoic chamber. A metamaterial aperture antenna was first developed as a microwave RRS to obtain the TSSRFs, consisting of many subwavelength, complementary electromagnetic inductive-capacitive (cELC) resonators, each cELC being engraved on a microstrip line and fed with a signal by a waveguide pattern. Since the geometry of the cELC controls the variation of its resonant
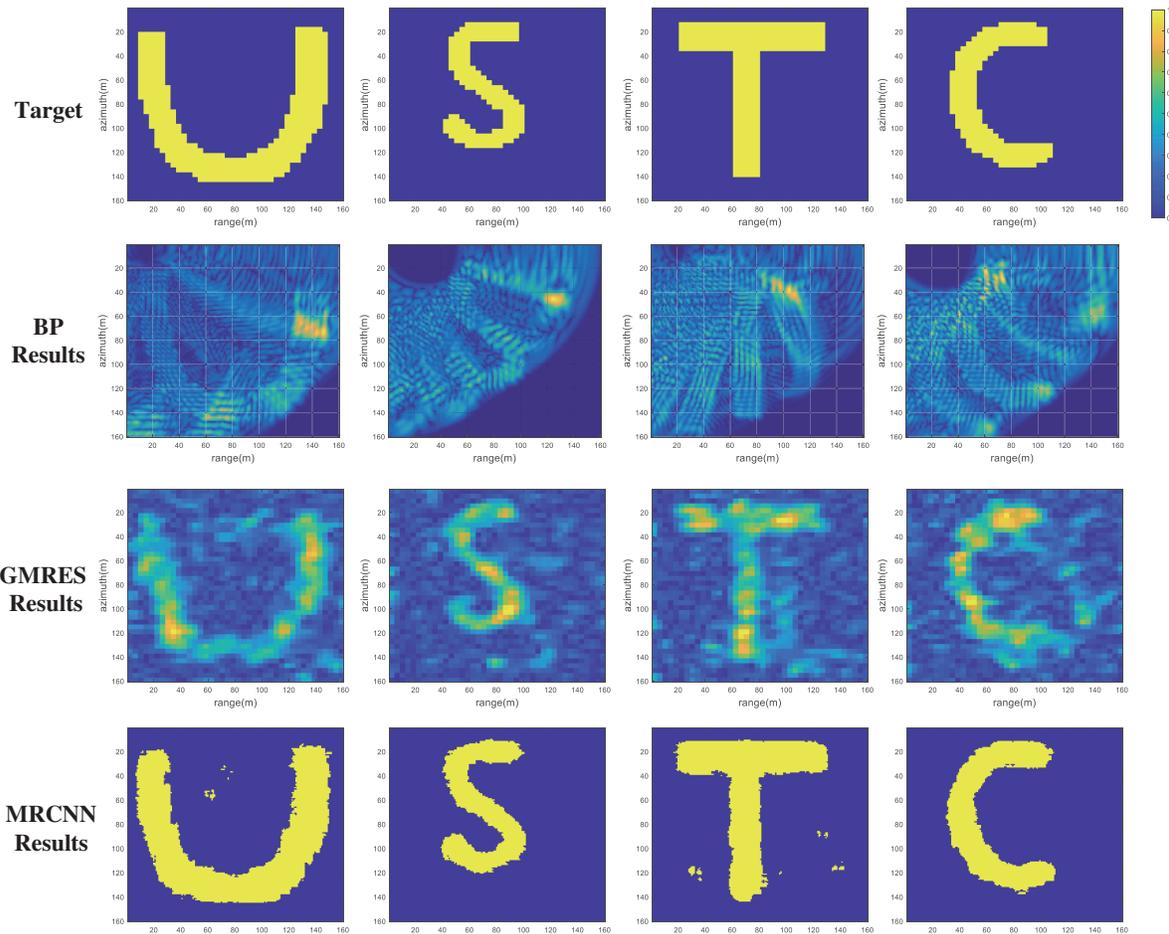
Figure 9. Reconstruction results of letter-shaped targets in the second, third and fourth rows for the BP algorithm, the GMRES algorithm in the traditional MSCI and the MRCNN, respectively. The first row shows the targets.
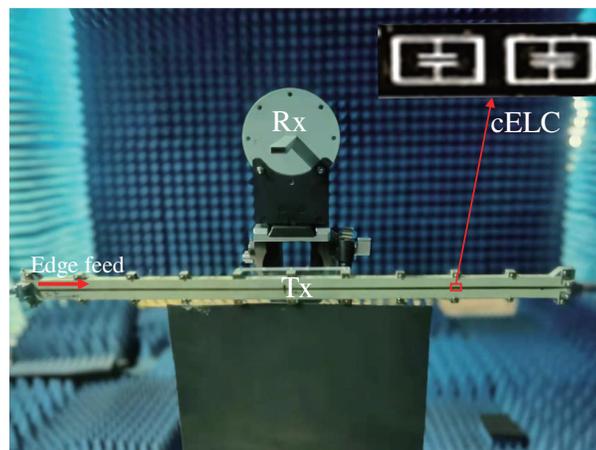
Figure 10. Experimental imaging configuration of the metamaterial aperture antenna.

frequency, it is possible to vary the geometric parameters of the resonator to obtain multiple cELCs with different resonant frequencies. 126 cELC resonators are etched on randomly distributed microstrip lines, containing 36 types of cELC resonators with different resonant frequencies, covering the K-band from 18 GHz to 22 GHz. When the resonant frequencies are swept in the available bandwidth, these cELCs form an aperture by leaking to generate a series of frequency-dependent varying TSSRFs. So the 1-D waveguide generates fan beams which vary only in azimuth. The top and bottom surfaces of the metamaterial aperture antenna are fabricated using perfect electric conductor (PEC) and the substrate is made of Rogers RO4003 with permittivity $= 3.55$ and thickness $= 1.58$ mm. The fabricated antennas are fed on one side by a standard SubMiniature version A (SMA) adapter, and the other side is connected to a $50\,\Omega$ matching load. The antenna array has microstrip transmission lines with width $= 3.2$ mm, length $= 600$ mm, and the dielectric plate and the ground with width $= 16$ mm, length $= 600$ mm.

The open waveguide receives signals of different TSSRF patterns at operating frequencies. The imaging experimental configuration is shown in Figure 10. The circuit connections in the imaging system are shown schematically in Figure 11. The two ports of the Vector Network Analyzer (VNA) are connected to the transmitting and receiving antennas, respectively, and the measured forward transmission coefficient (S_{21}) is the target scattering signal in the imaging scene. A simple sparse scene is constructed in an anechoic chamber of $4 \times 4 \times 3$ m, as shown in Figure 12, with the transmitting and receiving antennas at a height of 1.1 m and 1.3 m, respectively, and the absorbing materials placed in the gap between the transmitting and receiving antennas; three corner reflectors of 10 cm diameter were used as imaging targets, fixed on a slide rail 1 m distant in front of the antenna. Initially, a transmitting signal with a power of -30 dB is set on the VNA. The radar cross-section (RCS) of the corner reflector, viewed along the axis of symmetry of the line-of-sight (LOS), is:

$$\mathrm{RCS} = \frac{4\pi A_{\mathit{eff}}^2}{\lambda^2} \tag{13}$$

where A_{eff} is the effective area contributing to multiple internal reflections, and λ is the operating wavelength.
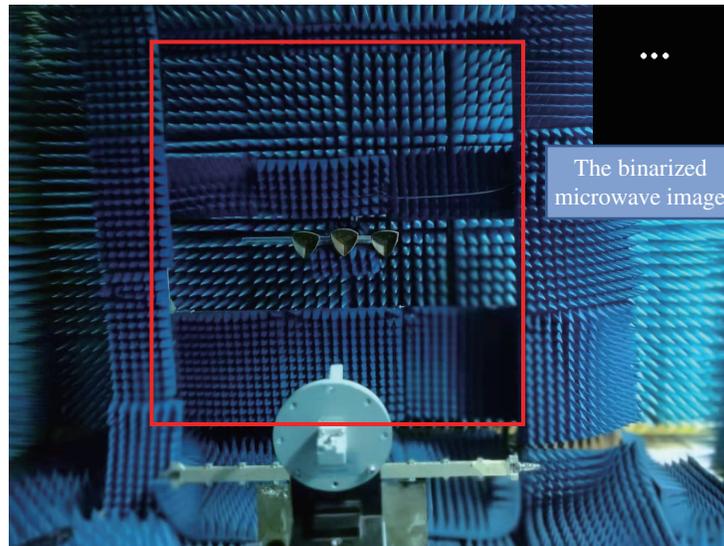
The backscattering coefficient is:

$$\sigma = \frac{RCS}{A} \tag{14}$$

where A is the radiation area of the antenna beam. The backscattering coefficient of the corner reflector is obtained to be approximately 10. Additional system parameters are given in Table 5.

After demonstrating the anechoic chamber imaging system, the experimental dataset for the intelligent MSCI needed to be fabricated. The designed 1-D metamaterial aperture antenna, which
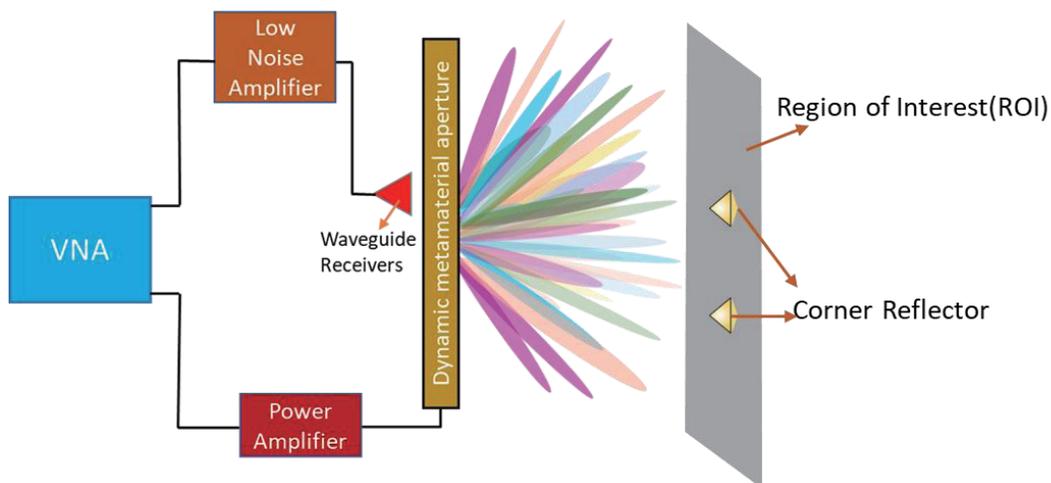
Figure 11. Schematic diagram of the circuit connection in the imaging system. The metamaterial aperture radiation waveform to the region of interest at a certain resonant frequency.

Figure 12. Imaging scene. The upper right corner shows its microwave image in the imaging area, with the brightness component of the image selected as the backscattering coefficient of the measured target.

Table 5. Experiment parameters.

Parameter	Value
Working frequency	18–22 GHz
Sampling interval	10 MHz
Number of orthogonal scenario sets	114
Number of radiation field types	400
Number of resonant units	126
Number of resonant unit types	36
Imaging area	$1.6\,\text{m} \times 1.6\,\text{m}$
cross-range resolution at 1 m	3.78 cm
Grid size	1 cm

produces TSSRFs in the operating band with undulating changes in the horizontal direction and gentling fluctuations in the vertical direction, can only be used for 1-D imaging [29–33]. When the 1-D metamaterial aperture antenna serves as an RRS, the imaging system can be considered monostatic imaging at a small area with limited transmitting perspective. Corner reflectors are chosen as imaging targets because they are suitable for monostatic sounding, and their backscatter coefficient is larger. The echoes from different targets are more weakly correlated, and it is easier for the neural network to extract information from the echoes to map the target image. A set of 114 mutually orthogonal target scenes was created by controlling the combination and position of corner reflectors in the horizontal direction on a sliding track. The operating metamaterial aperture antenna generates a series of frequency-dependent TSSRFs and radiates at each target scene. The frequency sweep interval is set to 10 MHz, which can ensure that the different TSSRFs of the metamaterial aperture antenna have a weak correlation from the simulation and measurement. A total of 400 different TSSRFs are generated at 400 frequencies, and S21s obtained at different TSSRF patterns are used as the input to the MRCNN. The backscattering coefficient distribution in the imaging region is used as the output of MRCNN, and then the samples are randomly grouped for training and testing the MRCNN.
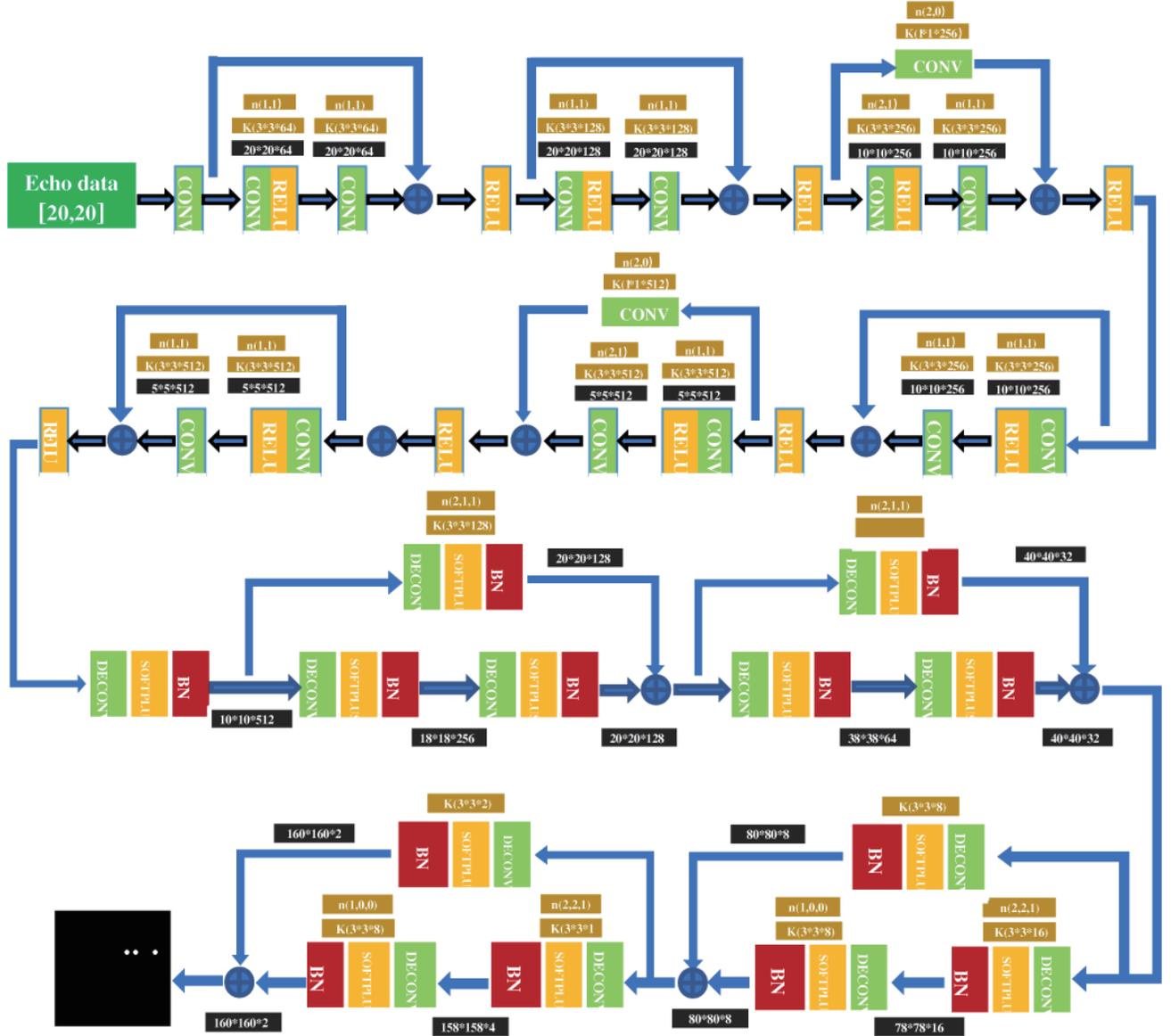
Figure 13. MRCNN for experimental scenes.

Due to the small amount of data, we again designed a specific network structure for this experimental imaging, as shown in Figure 13. In the network, a 3×3 convolution kernel has been used several times to increase the depth of the network without changing the data size. The convolution layer uses 6 residual modules with 13 convolutions, which convert the echo data of 20×20 into tensors of $5 \times 5 \times 512$, and the feature numbers are 1, 64, 128, 256, and 512, respectively, The deconvolution layer uses 4 residual modules with 8 deconvolutions, which convert from tensors of $5 \times 5 \times 512$ into tensors of $160 \times 160 \times 2$. Since the categories are target and background, the last deconvolution residual modules are used to change the number of channels to 2, which facilitates backpropagation to update the network parameters compared to the backscattering coefficients in the corresponding imaging area.

4.2. Experimental Results of Applying MRCNN

The 114 targets and corresponding echo data were randomly divided into two parts, where 108 groups are for training and the rest for testing. The gradient descent optimization algorithm with momentum
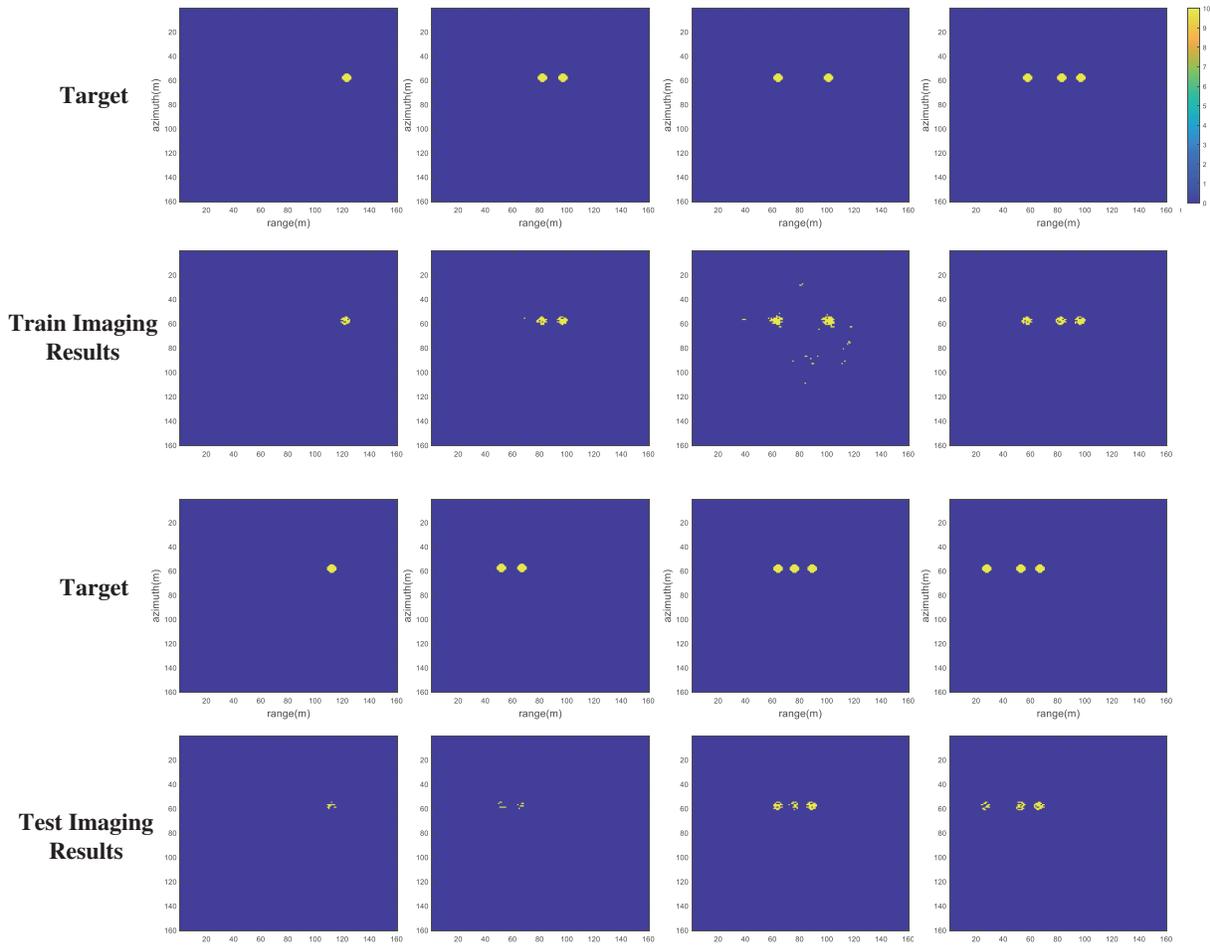
Figure 14. The output images obtained by MRCNN from the corresponding echo data of the training and test sets in the second and fourth rows respectively. The first and third rows show the 8 targets.

is used in the training, with a small batch size of 1 and the epoch set to 150. The learning rate is set to 10^{-3} and halved after every 50 epochs. The MRCNN is shown in Figure 13. The training of the network takes about 3 h. Four targets are randomly selected in each of the training and test sets, as shown in the first and third rows of Figure 14, and the output images by inputting the corresponding echo data into MRCNN in the training set and test set are shown in the second and fourth rows, respectively. It shows that MRCNN can accurately reconstruct the images of different locations and numbers of targets in the azimuthal region. Figure 15 represents the training loss and testing loss of MRCNN during 150 epochs, respectively. The NMSE results of the reconstructed images in the training and test sets are shown in Table 6.

Table 6. NMSE results for the reconstructions in Figure 15.

Column Sequence	1	2	3	4
Training Set	0.0085	0.0095	0.0104	0.0094
Testing Set	0.0126	0.0586	0.0098	0.0112

As shown in Table 6, the NMSE of the intelligent MSCI reconstructed target image is less than 0.1, and an important conclusion can be drawn: Without changing the imaging scene, the echoes are obtained simply by changing the position of the corner reflector in the same azimuthal region. With the trained
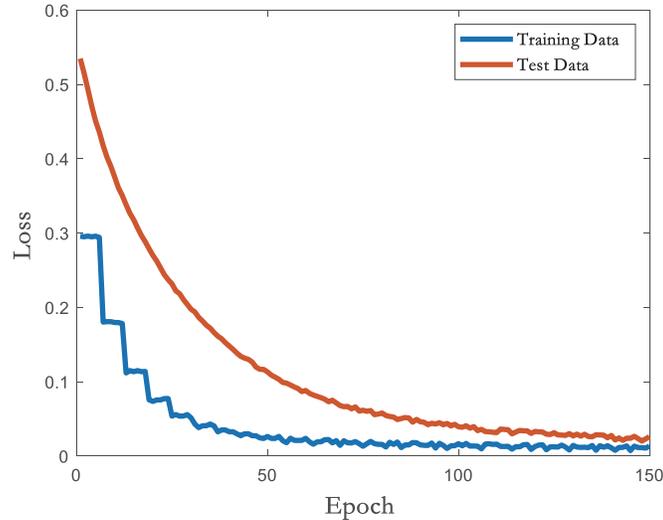
Figure 15. The training loss and test loss of the MRCNN for experimental scenes during 150 epochs.

MRCNN mapping, the MRCNN can directly show the image of the corner reflector at an unknown position and number in the imaging region. This demonstrates that MRCNN has learned the intelligent mapping model of the MSCI system in the same MSCI scenario. The essential reason that MRCNN can provide better image results is that the present neural network, consisting of 15 convolutional processing layers and 13 deconvolutional processing layers (a total of 1.1 million parameters), provides approximate numerical solutions to the imaging equations in MSCI.

5. CONCLUSION

This paper proposes an intelligent MSCI method based on the DNN and establishes a fundamental link between MSCI and DNN, and then develops a multi-level residual convolutional neural network framework for MSCI. MRCNN utilizes the strong feature extraction capability of the convolutional layer network and the target feature mapping imaging capability of the deconvolutional layer network and adds the residual modules so that the neural network model becomes easier to optimize, which can construct an end-to-end mapping between scattering echo and target imaging after autonomous learning. To verify the effectiveness of the intelligent MSCI method, the simulation is compared with the GMRES optimization algorithm in traditional MSCI. To test the generalization ability of the intelligent MSCI method, MRCNN is trained on the MNIST dataset, and the test targets are composed of letters with the backscattering coefficient of 1. MRCNN can still reconstruct the high-resolution image within 1 s. Meanwhile, a metamaterial aperture antenna is developed as the RRS; the corner reflector is used as the target; the combination and position of the corner reflector are changed in the horizontal direction to produce a number of mutually orthogonal target scenes. The acquired data is used for training and testing MRCNN. The experiments show that the trained MRCNN can still show the image of the corner reflector at an unknown position and number in the imaging region. More advanced network architectures can reconstruct more practical and complex targets, which are to be explored in our further studies. When complex scenes are imaged, it is also necessary to develop an effective RRS to acquire a dataset. The intelligent MSCI method changes the traditional MSCI strategy for solving the TSSRFs to reconstruct imaging, providing an ideal solution for MSCI.

ACKNOWLEDGMENT

REFERENCES

1. Dunkel, R., R. Saddler, and A. Doerry, "Synthetic aperture radar for disaster monitoring," *Radar Sensor Technology XV*, 125–134, 2011.

2. Madsen, S., W. Edelstein, L. D. DiDomenico, and J. LaBrecque, "A geosynchronous synthetic aperture radar; for tectonic mapping, disaster management and measurements of vegetation and soil moisture," *IEEE 2001 International Geoscience and Remote Sensing Symposium (IGARSS)*, Vol. 1, 447–449, 2001.

3. Deng, B., Y. Qin, Y. Li, H. Wang, and X. Li, "A novel approach to range doppler SAR processing based on legendre orthogonal polynomials," *IEEE Geoscience and Remote Sensing Letters*, Vol. 6, No. 1, 13–17, 2008.

4. Zhu, Y., Y. Su, and W. Yu, "An ISAR imaging method based on MIMO technique," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48, No. 8, 3290–3299, 2010.

5. Guo, Y., D. Wang, X. He, and B. Liu, "Super-resolution staring imaging radar based on stochastic radiation fields," *2012 IEEE MTT-S International Microwave Workshop Series on Millimeter Wave Wireless Technology and Applications*, 1–4, 2012.

6. Li, D., X. Li, Y. Qin, Y. Cheng, and H. Wang, "Radar coincidence imaging: An instantaneous imaging technique with stochastic signals," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 52, No. 4, 2261–2277, 2013.

7. Xu, X., X. Zhou, Y. Cheng, and Y. Qin, "Radar coincidence imaging with array position error," *2015 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, 1–4, 2015.

8. Zhou, X., H. Wang, Y. Cheng, and Y. Qin, "Sparse auto-calibration for radar coincidence imaging with gain-phase errors," *Multidisciplinary Digital Publishing Institute*, Vol. 15, No. 11, 27611–27624, 2015.

9. Zhou, X., H. Wang, and Y. Cheng, "Expansion-compression variance-component-based autofocusing method for joint radar coincidence imaging and gain-phase error calibration," *SPIE*, Vol. 11, No. 2, 025002, 2017.

10. Cao, K., X. Zhou, Y. Cheng, and Y. Qin, "Improved focal underdetermined system solver method for radar coincidence imaging with model mismatch," *Journal of Electronic Imaging*, Vol. 26, No. 3, 033001, 2017.

11. Cao, K., X. Zhou, Y. Cheng, B. Fan, and Y. Qin, "Total variation-based method for radar coincidence imaging with model mismatch for extended target," *Journal of Electronic Imaging*, Vol. 26, No. 6, 063007, 2017.

12. He, X., B. Liu, S. Chai, and D. Wang, "A novel approach of high spatial-resolution microwave staring imaging," *Conference Proceedings of 2013 Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, 75–78, 2013.

13. Meng, Q., B. Liu, C. Tian, Y. Guo, and D. Wang, "Correlation algorithm of microwave staring correlated imaging based on multigrid and CGLS," *2015 IEEE International Conference on Communication Problem-Solving (ICCP)*, 359–362, 2015.

14. Yu, H., G. H. Lu, and H. L. Zhang, "Adaptive sparse recovery of moving targets for distributed MIMO radar," *Advanced Materials Research*, Vol. 933, 450–455, Trans. Tech. Publ., 2014.

15. Lou, G. and H. Shi, "Face image recognition based on convolutional neural network," *China Communications*, Vol. 17, No. 2, 117–124, 2020.

16. Shen, W. and W. Wang, "Node identification in wireless network based on convolutional neural network," *2018 14th International Conference on Computational Intelligence and Security (CIS)*, 238–241, 2018.

17. Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in Neural Information Processing Systems*, Vol. 27, 2014.

18. Ledig, C., L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al., "Photo-realistic single image super-resolution using a generative adversarial network," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4681–4690, 2017.

19. Long, J., E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3431–3440, 2015.

20. Girshick, R., J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 580–587, 2014.

21. Milletari, F., N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," *2016 Fourth International Conference on 3D Vision (3DV)*, 565–571, 2016.

22. Niu, S., X. Qiu, L. Peng, and B. Lei, "Parameter prediction method of SAR target simulation based on convolutional neural networks," *EUSAR 2018; 12th European Conference on Synthetic Aperture Radar*, 1–5, 2018.

23. Shao, W. and Y. Du, "Microwave imaging by deep learning network: Feasibility and training method," *IEEE Transactions on Antennas and Propagation*, Vol. 68, No. 7, 5626–5635, 2020.

24. Li, L., Y. Shuang, Q. Ma, H. Li, H. Zhao, M. Wei, C. Liu, C. Hao, C.-W. Qiu, and T. J. Cui, "Intelligent metasurface imager and recognizer," *Light: Science & Applications*, Vol. 8, No. 1, 1–9, 2019.

25. Li, H.-Y., H.-T. Zhao, M.-L. Wei, H.-X. Ruan, Y. Shuang, T. J. Cui, P. Del Hougne, and L. Li, "Intelligent electromagnetic sensing with learnable data acquisition and processing," *Patterns*, Vol. 1, No. 1, 100006, 2020.

26. Xing, Y., L. Zhong, and X. Zhong, "An encoder-decoder network based FCN architecture for semantic segmentation," *Wireless Communications and Mobile Computing*, Vol. 2020, 2020.

27. Deng, L., "The mnist database of handwritten digit images for machine learning research [best of the web]," *IEEE Signal Processing Magazine*, Vol. 29, No. 6, 141–142, 2012.

28. He, K., X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778, 2016.

29. Hunt, J., T. Driscoll, A. Mrozack, G. Lipworth, M. Reynolds, D. Brady, and D. R. Smith, "Metamaterial apertures for computational imaging," *Science*, Vol. 339, No. 6117, 310–313, 2013.

30. Sleasman, T., M. Boyarsky, M. F. Imani, J. N. Gollub, and D. R. Smith, "Design considerations for a dynamic metamaterial aperture for computational imaging at microwave frequencies," *JOSA B*, Vol. 33, No. 6, 1098–1111, 2016.

31. Imani, M. F., T. Sleasman, and D. R. Smith, "Two-dimensional dynamic metasurface apertures for computational microwave imaging," *IEEE Antennas and Wireless Propagation Letters*, Vol. 17, No. 12, 2299–2303, 2018.

32. Hoang, T. V., V. Fusco, T. Fromenteze, and O. Yurduseven, "Computational polarimetric imaging using two-dimensional dynamic metasurface apertures," *IEEE Open Journal of Antennas and Propagation*, Vol. 2, 488–497, 2021.

33. Sleasman, T. A., M. F. Imani, A. V. Diebold, M. Boyarsky, K. P. Trofatter, and D. R. Smith, "Implementation and characterization of a two-dimensional printed circuit dynamic metasurface aperture for computational microwave imaging," *IEEE Transactions on Antennas and Propagation*, Vol. 69, No. 4, 2151–2164, 2020.